

Fractal Analysis of Time Series and Distribution Properties of Hurst Exponent

Malhar Kale[†]
Ferry Butar Butar[‡]

Abstract

Fractal analysis is done by conducting rescaled range (R/S) analysis of time series. The Hurst exponent and the fractal (fractional) dimension of a time series can be estimated with the help of R/S analysis. The Hurst exponent can classify a given time series in terms of whether it is a random, a persistent, or an anti-persistent process. Simulation study is run to study the distribution properties of the Hurst exponent using first-order autoregressive process. If time series data are randomly generated from a normal distribution then the estimated Hurst Exponents are also normally distributed.

Fractals: Concepts and Analysis

Introduction

The term fractal was coined by Benot Mandelbrot in 1975, from the Latin fractus or “fraction/ broken”. The concept of fractals is that they have a large degree of self similarity within themselves. They are copies of themselves buried deep within the original. They also reveal infinite detail.

Definition 1. A *Fractal* is “a geometric shape that is self-similar and has fractional dimensions”.

Peters (1994) stated, “Fractal geometry is the geometry of the Demiurge. Unlike Euclidean geometry, it thrives on roughness and symmetry”. Objects are infinitely complex. The more closely one examines them, the more detail they reveal. A fir tree, for example, is a fractal form. It can be visualized as a cone sitting atop a rectangle. Yet, fir trees are not cones and rectangles. They are a complex network of branches, no branch being the same as any other. According to Peters (1994), “Euclidean geometry cannot replicate a tree. Euclidean geometry recreates the perceived symmetry of the tree, but not the variety that actually builds its structure. Underlying this perceived symmetry is a controlled randomness. This “self-similar” quality is the defining characteristic of fractals”.

Euclidean geometry has a primary characteristic that dimensions are integers. For examples, lines are one-dimensional, planes are two-dimensional, and solids are three dimensional. In general, Euclidean shapes are smooth, continuous, and homogeneous. Consider a spherical sponge ball. Though it resides in a three dimensional space, it has irregularities or holes. It is not smooth, continuous, or homogeneous. So it would be fair enough to say that, dimensionally, it is something more than a plane, but less than a solid. Its dimension is somewhere between two and three. It has a non-integer dimension

or a fractional dimension. It is a fractal. Similarly, a Sierpinski triangle clearly has higher dimensionality than a line, but less dimensionality than a plane. It has triangle shaped holes in it. Its dimension is between one and two. Offering the same explanation, we see that a time series is a fractal. A fractal dimension defines how an object fills the space it occupies.

The fractal dimension of a time series measures how serrated or jagged the given time series is. As we would expect, a straight line has a fractal dimension of 1. It is the same as its Euclidean dimension. A time series following a Gaussian random walk has a fractal dimension of 1.5. If the fractal dimension of a time series is between 1 and 1.5, the time series is somewhere between a straight line and Gaussian random walk. The importance of the fractal dimension of a time series lies in the fact that it recognizes that a process can be somewhere between deterministic and random. See Peters (1996).

Measuring the Fractal Dimension

Fractal dimension of a time series is calculated by measuring how jagged it is. We count the number of circles of a given, fixed diameter that are needed to cover the entire time series. We then increase the diameter and again count the number of circles required to cover the time series. If we continue this process we see that, the number of circles has an exponential relationship with the radius of the circle. The number of circles is related to the diameter of the circle according to the following relationship:

$$N * d^D = 1 \quad (1)$$

where, N = number of circles, d = diameter, and D = fractal dimension.

Equation (1) can be transformed to find the fractal dimension:

$$D = \frac{\log(N)}{\log\left(\frac{1}{d}\right)} \quad (2)$$

The Hurst Exponent

The Hurst Exponent is directly related to the fractal dimension, which measures the smoothness of a surface, or, in our case, the smoothness of a time series. The relationship between the fractal dimension D , and the Hurst Exponent H , is given by:

$$D = 2 - H \quad (3)$$

where, $0 \leq H \leq 1$. The closer the value of the Hurst Exponent to 0, the more jagged will the time series be.

Definition 2. The *Hurst Exponent* is the measure of the smoothness of fractal time series based on the asymptotic behavior of the rescaled range of the process.

The Hurst Exponent, H , can be estimated by:

$$H = \frac{\log(R/S)}{\log(T)} \quad (4)$$

where T is the duration of the sample data and R/S the corresponding value of the rescaled range.

Definition 3. *Rescaled Range* is the “measure characterizing the divergence of time series defined as the range of the mean-centered values for a given duration (T) divided by the standard deviation for that duration”.

Hurst (1965) proposed:

$$R/S = k * T^H \quad (5)$$

where k is a constant that depends on the time series. See Gleick (1988), Peters (1994, 1996).

Rescaled Range Analysis

Rescaled Range

Hurst (1965) developed the rescaled range analysis, a statistical method to analyze long records of natural phenomenon. Rescaled Range Analysis is the central tool of fractal data modeling. The two factors used in this range analysis are: 1) The difference between the maximum and the minimum cumulative values, and 2) The standard deviation from the observed values.

The R/S value in equation (5) is known as the rescaled range because it has a mean of zero and is expressed in terms of local standard deviation. The R/S value scales as we increase the time increment, T , by a power law value which equals H , the Hurst Exponent. All fractals scale according to a power law. By rescaling data, we can compare diverse phenomena and time periods. Rescaled Range Analysis enables us to describe a time series that has no characteristic scale.

Brownian motion is the primary model for a random walk process. Einstein (1908) found the distance a particle covers increases with respect to time according to the following relation:

$$R = T^{0.5} \quad (6)$$

where R is the distance covered by the particle in time T . See Peters (1994, 1996).

We begin with a time series F_1, \dots, F_N . The mean value, μ_N and the standard deviation, S_N , of the time series are estimated. The rescaled range is calculated by rescaling the time series by subtracting the sample mean:

$$S_i = F_i - \hat{\mu}_N; \quad i = 1, \dots, N \quad (7)$$

The resulting series has a mean of zero. Now we form a cumulative time series Γ :

$$\Gamma_i = S_1 + S_i; \quad i = 2, \dots, N \quad (8)$$

The adjusted range is the difference between the maximum value and the minimum value of Γ_i :

$$R_N = \max(\Gamma_1, \dots, \Gamma_N) - \min(\Gamma_1, \dots, \Gamma_N) \quad (9)$$

The maximum value of Γ_i will always be greater than or equal to zero, and, the minimum value will always be less than or equal to zero. Therefore the adjusted range will always be non-negative. The adjusted range R_N is the distance that the system travels for time N . We can use equation (6) only if the time series we are considering is independent of increasing values of N . To take into account the fact that economic time series systems are not independent with respect to time, Hurst (1965) found a more general form of equation (6):

$$(R/S)_N = c * N^H, \quad (10)$$

where the subscript N for $(R/S)_N$ refers to the R/S value for F_1, \dots, F_N ; c is a constant, and H is the Hurst Exponent.

Estimation of the Hurst Exponent

We will now take a look at the estimation of Hurst Exponent using the Rescaled Range Analysis. The estimation procedure involves three basic steps.

Step 1: The cumulative total at each point in time, for a time series over a total duration N , is given by:

$$\Gamma_{N,k} = \sum_{i=1}^k (F_i - \mu_N), \text{ for } 0 < k \leq N, \quad (11)$$

where F_i = the value of the time series at time i , μ_N = the mean over the whole data set given by:

$$\mu_N = \frac{1}{N} \sum_{i=1}^N F_i \quad (12)$$

The range R of Γ , is given by:

$$R = \text{Max}(\Gamma_{N,k}) - \text{Min}(\Gamma_{N,k}), \quad (13)$$

where $\text{Max}(\Gamma_{N,k})$ = the maximum value of $\Gamma_{N,k}$ and $\text{Min}(\Gamma_{N,k})$ = the minimum value of $\Gamma_{N,k}$. See Peters (1996).

The standard deviation of the values over the whole data set is given by:

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [F_i - \mu_N]^2} \quad (14)$$

The Rescaled Range is given by: R/S

Step 2: We now take $N = N/2$. The Rescaled Range R/S is calculated following the procedure shown in Step 1 for the two segments. The average value of R/S is then calculated. We repeat the entire procedure for successively smaller intervals over the data set, dividing each segment obtained in each step in two and calculating R/S for each segment and finding the average R/S . Pictorially the R/S calculation looks like figure 1. See Saupe (1988).

Step 3: The Hurst exponent is estimated by plotting the values of $\log(R/S)$ versus $\log(N)$. The slope of the best fitting line gives the estimate of the Hurst exponent. This is done through an ordinary least squares regression procedure. We can generate “least squares” regression line using linear regression method.

Figure 1: Rescaled Range Analysis Process for estimating Hurst Exponent



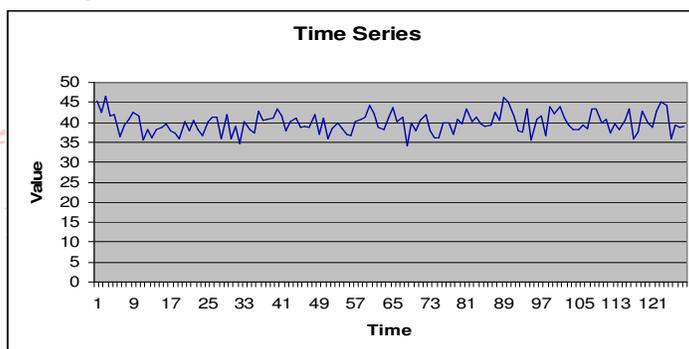
Inferences from the Hurst Exponent

The value of Hurst Exponent varies between 0 and 1. $H = 0.5$ implies a random walk or an independent process. If $0 \leq H < 0.5$ then we have anti-persistence. Such a process covers less “distance” than a random walk. This means that a decreasing process or a time series, then, it is more probable that the process or the time series will show an increasing trend. An anti-persistent time series will exhibit higher noise and more volatility. If $0.5 < H \leq 1$ then we have persistence. A persistent process or a time series will cover more “distance” than a random walk. This means that a decreasing process or a time series, then, it is more probable that the process will continue to decrease, and if we have an increasing time series, then, it is more probable that the time series will continue to show an increasing trend. A persistent time series has long memory effects. In theory, the trend at a particular point in time affects the remainder of the time series. A persistent time series will exhibit higher noise and more volatility.

Definition 4. *Volatility* is “the statistical measure of the tendency of time series to rise or fall sharply within a period of time”.

Example 1. The time series shown next (Figure 2) was taken for the purpose of estimating the Hurst Exponent.

Figure 2: Time Series collected over 128 data values

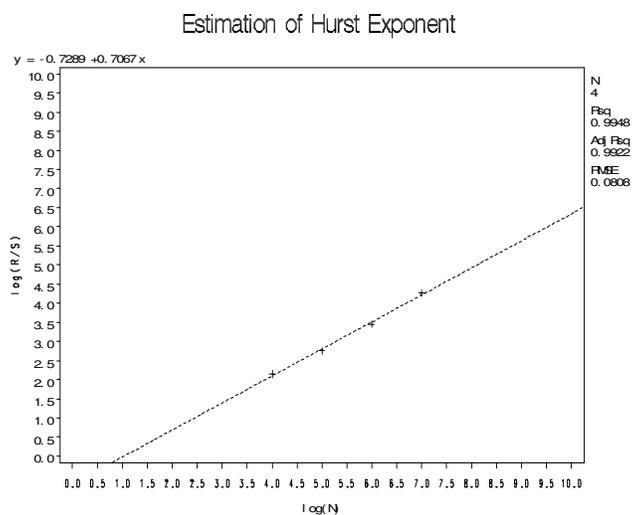


The data were collected over 128 time intervals and the rescaled range analysis gave the values in Table 1 shown below:

Table 1. Rescaled Range Analysis values to estimate H

Segment size (N)	$\log_2(N)$	$\log_2(R/S)$
128	7	4.2754
64	6	3.4539
32	5	2.7479
16	4	2.1550

Figure 3: Regression Line which estimates the Hurst Exponent



The slope of the log – log plot of R/S versus N gives the estimate of the Hurst exponent. SAS output gave the estimate of H to be 0.7067. See Cody and Smith (1997).

Time Series Models and the First Order Autoregressive Process

A time series is a set of observation x_i 's, each one being recorded at a specific time i . Time series can be of a discrete nature or a continuous nature. A discrete-time time series is one in which the set of times when the observations are recorded is a discrete set. A continuous-time time series is one in which the observations are recorded continuously over some time interval.

Definition 5. A *Time Series Model* for the observed data $\{x_t\}$ is “a specification of joint distributions of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization”. See Brockwell and Davis (1996).

A Zero-Mean Model: IID Noise

In one of the simplest of the models for a time series, there is no trend or a seasonal component and the observations are independent and identically distributed random variables with zero mean. Such a sequence of random variables X_1, \dots, X_n is referred to as IID noise. By definition, for any positive integer n and real numbers x_1, \dots, x_n ,

$$\begin{aligned} P[X_1 \leq x_1, \dots, X_n \leq x_n] &= P[X_1 \leq x_1] \dots P[X_n \leq x_n] \\ &= F(x_1) \dots F(x_n) \end{aligned}$$

where, $F(\cdot)$ is the cumulative distribution function of each of the IID variables X_1, \dots, X_n . There is no dependence between observations in the IID noise model. For all $h \geq 1$ and for all x, x_1, \dots, x_n ,

$$P[X_{n+h} \leq x | X_1 = x_1, \dots, X_n = x_n] = P[X_{n+h} \leq x].$$

This shows that the knowledge of random variables X_1, \dots, X_n cannot be used for predicting the behavior of X_{n+h} .

Stationary Models and White Noise

A time series $\{X_t\}$ is said to be stationary if it has statistical properties similar to those of the time-shifted series $\{X_{t+h}\}$, for each integer h .

Definition 6. The *Mean Function* of time series $\{X_t\}$ with $E(X_t^2) \leq \infty$ is:
 $\mu_X(t) = E(X_t)$

Definition 7. The *Covariance Function* of $\{X_t\}$ is:

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

for all integers r and s .

Definition 8. A time series $\{X_t\}$ is (Weakly) *Stationary* if $\mu_X(t)$ is independent of t , and $\gamma_X(t+h, t)$ is independent of t for each h .

Definition 9. If $\{X_t\}$ is a sequence of uncorrelated random variables, each with zero mean and finite variance σ^2 , then it can be clearly seen that $\{X_t\}$ is stationary with same covariance function as IID noise. Such a sequence is referred to as white noise with mean 0 and variance σ^2 , and, is denoted by:

$$\{X_t\} \sim \text{WN}(0, \sigma^2)$$

First Order Autoregressive or AR(1) Process

We assume that $\{X_t\}$ is a stationary series satisfying the following equation:

$$X_t = \phi X_{t-1} + Z_t, \quad t = \pm 1, \dots, \quad (15)$$

where $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, $|\phi| < 1$ and $\{Z_t\}$ is uncorrelated with X_s for each $s < t$. An AR(1) process is defined by the stationary solution $\{X_t\}$ of equation (17). Here ϕ is the autoregressive coefficient. The initial observation X_0 can be treated as fixed or stochastic. When X_0 is fixed with $|\phi| < 1$, the model is stationary asymptotically. When X_0 is stochastic or random, the model is stationary. If $|\phi| = 1$ then we have a random walk model. If $|\phi| = 0$ then we have white noise. That is,

$$X_t = Z_t, \quad t = \pm 1, \dots,$$

Z_t is drawn from a distribution with mean 0 and a finite variance. If $|\phi| > 1$, this would mean the model would be explosively increasing or decreasing.

Tests for Normality

Goodness-of-Fit Tests for Normality

Goodness-of-Fit tests are conducted to see whether a given data come from a population with a specific distribution. Goodness-of-Fit tests are of the form: H_o = Data follows a specific distribution and H_a = Data does not follow a specific distribution. To test for normality of data we can conduct the following tests:

Shapiro-Wilk Test

Test statistic W is calculated to test whether a random sample comes from a normal distribution. The Shapiro-Wilk's W can be calculated as follows:

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

where, the $x_{(i)}$ s are the ordered sample values and the a_i s are constants. The value of $\text{Prob} < W$ is the p-value. The p-value is compared with the chosen alpha (α) level. If the p-value is less than α then the null hypothesis that the data are normally distributed is rejected. Otherwise, the null hypothesis is not rejected and we have sufficient evidence to suggest that the data come from a normally distributed population.

Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test is based on the empirical distribution function (ECDF). Given N ordered points x_1, \dots, x_N the ECDF is defined as:

$$E_N = \frac{n(i)}{N}$$

where, $n(i)$ is the number of points less than x_i and the x_i are ordered from the smallest to the largest value. The Kolmogorov-Smirnov statistic is defined as:

$$D = \max_{1 \leq i \leq N} \left| F(x_i) - \frac{n(i)}{N} \right|$$

where, F is the theoretical cumulative distribution of the continuous distribution being tested. In our case the distribution is normal. The null hypothesis is rejected if the test statistic D is greater than the critical value obtained from the table. The value of $\text{Prob} > D$ is the p-value. If the p-value is less than α then the null hypothesis that the data are normally distributed is rejected.

Anderson-Darling Test

Anderson-Darling test statistic is defined as:

$$A^2 = -N - S \quad \text{where,}$$
$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(x_i) + \ln(1 - F(x_{N+1-i}))]$$

Here, F is the cumulative distribution function of the distribution being tested and x_i are ordered. The test is one sided and we reject the null hypothesis if the value of A is greater than the critical value. Also, the value of $\text{Prob} > A^2$ is the

p-value. If the p-value is less than the level of significance α then the null hypothesis is rejected.

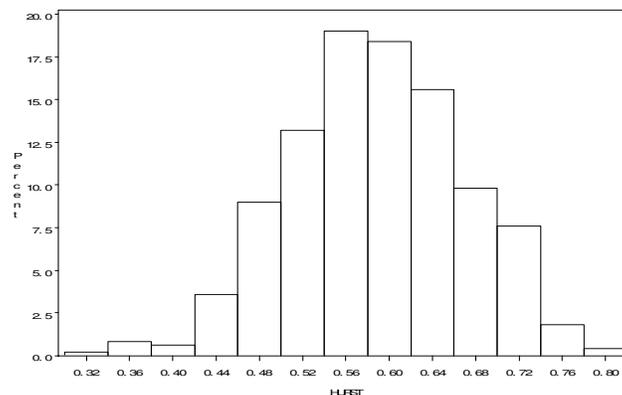
Distribution Properties of Hurst Exponent

In our study we use model (15) by letting $\phi = 0$ and we draw Z_t from a normal distribution with mean 0 and variance 1. To determine the distribution properties of the Hurst exponent, we simulate 128 values of random data for a time series. For each data set, the Hurst exponent H is estimated using R/S analysis. 500 values of the Hurst exponent were recorded. The random data were generated using the SAS `rannor()` function. The function generates a randomly chosen value for a normally distributed variable with population mean of 0 and standard deviation of 1.

The Capability Procedure `Proc Capability` (Cody and Smith, 1997) applied to the data set of the values of the Hurst exponent produces the Shapiro-Wilk, Kolmogorov-Smirnov, and the Anderson-Darling tests for normality. The p-values of the tests were recorded. Also, the Histogram and the Normal Probability plot were generated to test for normality of the data. Both of these techniques are visual assessments of normality, rather than formal tests of normality. The results are as follows:

Histogram

Figure 4: Histogram
TEST FOR NORMALITY

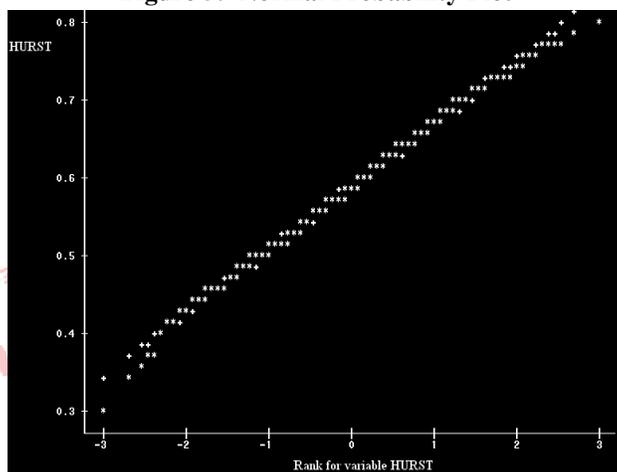


The histogram in Figure (4) suggests that the Hurst exponent is normally distributed.

Normal Probability Plot

Figure (5) is the Normal Probability plot. The plotted points show a strong linear pattern suggesting that the Hurst exponent is normally distributed.

Figure 5: Normal Probability Plot



Goodness-of-Fit Tests

Table 2 shows the values of different statistics for testing normality. The table is obtained from SAS output.

Table 2. Test for Normality, Test Statistic, and p-Value.

Test Statistic	Value	p-value
Shapiro-Wilk W	0.997	0.404
Kolmogorov-Smirnov D	0.021	0.150
Anderson-Darling A-Sq	0.243	0.250

The p-values of the Shapiro-Wilk, Kolmogorov-Smirnov, and the Anderson-Darling Tests are greater than the significance level $\alpha = 0.05$. So we don't have sufficient evidence to reject the null hypothesis and we can say that the data are normally distributed.

Summary

To study the distribution properties of Hurst Exponent simulations were conducted to randomly generate data values of several Time Series from a normal distribution. Rescaled Range Analysis performed on each of the generated Time Series estimated the values of Hurst Exponents. Tests for normality were conducted to check if the estimated values of Hurst Exponents were also normally distributed. The tests included Goodness-of-Fit Tests and visual tools such as Histograms and Normal Probability Plots. It was observed that if Time Series are randomly generated from a normal distribution, then, the estimated Hurst Exponents are also normally distributed. Future work will involve the study of distribution properties of Hurst Exponent for a random walk model for a Time Series. That is, a stationary series $\{X_t\}$ satisfying:

$$X_t = X_{t-1} + Z_t, \quad t = \pm 1, \dots,$$

Also, distribution properties of Hurst Exponent for a stationary model will be studied. A stationary series $\{X_t\}$ satisfies:

$$X_t = \phi X_{t-1} + Z_t, \quad t = \pm 1, \dots, \quad \text{where } |\phi| < 1.$$

† *Malhar kale*, Kaiser Permanente.

‡ *Ferry Butar Butar*, PhD, Sam Houston State University, USA

Journal Of

References

- Bowerman, L. B., & O'Connell, T. R. (2000). *Linear Statistical Models: An Applied Approach* (2nd ed.), Duxbury Press.
- Brockwell, J. P., & Davis, A. R. (1996). *Introduction to Time Series and Forecasting*, Springer-Verlag New York Inc.
- Cody, P. R., & Smith, K. J. (1997), *Applied Statistics and the SAS Programming Language* (4th ed.), Prentice-Hall, Inc..
- Freund, J. R., & Wilson, J. W. (1996). *Statistical Methods* (2nd ed.), Academic Press.
- Gleick, J. (1988). *Chaos: Making a New Science*, Penguin Books, J.
- Kutner, H. M., Nachtsheim, J. C., Neter, J., & Li, W. (1996). *Applied Linear Statistical Models*: (4th ed.), Irwin Professional Pub.
- Peters, E. E. (1996). *Chaos and Order in the Capital Markets* (2nd ed.), John Wiley & Sons, Inc. Peters, E. E. (1994). *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*, John Wiley & Sons, Inc.
- Ravi, A., & Butar Butar, F. (2010). *An insight into heavy-tailed distribution*, Mathematical Sciences & Mathematics Education, Vol. 5, no 1 pp. 11 – 25.
- Saupe, D. (1988). *Algorithms for Random Fractals*, Springer-Verlag New York Inc.